

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/132268/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Iliescu, Adela F. ORCID: <https://orcid.org/0000-0001-6642-9385>, Dwyer, Dominic M. ORCID: <https://orcid.org/0000-0001-8069-5508> and Honey, Robert C. ORCID: <https://orcid.org/0000-0001-6870-1880> 2020. Individual differences in the nature of conditioned behavior across a conditioned stimulus: adaptation and application of a model. *Journal of Experimental Psychology: Animal Learning and Cognition* 46 (4) , pp. 460-269. 10.1037/xan0000270 file

Publishers page: <http://dx.doi.org/10.1037/xan0000270>  
<<http://dx.doi.org/10.1037/xan0000270>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Individual differences in the nature of conditioned behavior across a conditioned stimulus:

Adaptation and application of a model

Adela F. Iliescu, Dominic M. Dwyer, & Robert C. Honey

Cardiff University

Short title: HeiDI adaptation

Accepted in May 2020 to:

*Journal of Experimental Psychology: Animal Learning and Cognition*

Address for correspondence: R. C. Honey, School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff, CF10 3AT, UK; Tel: +44 (0)29 20875868; Email: [Honey@cardiff.ac.uk](mailto:Honey@cardiff.ac.uk)

The development of HeiDI was supported by a grant awarded to R.C.H. and D.M.D. by the BBSRC (UK; BB/T004339/1), and a School of Psychology Studentship awarded to A.F.I.

© 2020, American Psychological Association. This paper is not the copy of record and may not exactly replicate the final, authoritative version of the article. Please do not copy or cite without authors' permission. The final article will be available, upon publication, via its DOI: 10.1037/xan0000270



### Abstract

Pavlovian conditioning procedures produce marked individual differences in the form of conditioned behavior. For example, when rats are given conditioning trials in which the temporary insertion of a lever into an operant chamber (the conditioned stimulus, CS) is paired with the delivery of food (the unconditioned stimulus, US), they exhibit knowledge of the lever-food relationship in different ways. For some rats (known as sign-trackers) interactions with the lever dominate, while for others (goal-trackers) approaching the food well dominates. A formal model of Pavlovian conditioning (HeiDI) attributes such individual differences in behavior to variations in the perceived salience of the CS and US. An application of the model in which the perceived salience of the CS declines (i.e., adapts) across its duration, predicts changes in these individual differences within the presentation of the CS: The sign-tracking bias is predicted to decline and goal-tracking bias is predicted to increase across the presentation of a lever. The accuracy of these predictions was confirmed through analysis of archival data from female and male rats.

*Keywords:* Rat, sign-tracking, goal-tracking, HeiDI, inhibition of delay

Formal models of Pavlovian conditioning in non-human animals have traditionally given no consideration to individual differences in conditioned behavior (e.g., Rescorla & Wagner, 1972; Mackintosh, 1975; Pearce & Hall, 1980; Stout & Miller, 2007; Wagner, 1981). This is remiss: Even simple conditioning procedures, which remain the principal test bed for such models, produce marked individual differences in the strength and nature of conditioned behavior. For example, consider a rat that has been placed in a chamber and every now and then a retractable lever is inserted into the chamber (the conditioned stimulus, CS) and upon its retraction a food pellet is delivered to an adjacent food well (the unconditioned stimulus, US). One conditioned behavior that develops in such a procedure is called goal-tracking (e.g., Boakes, 1977), where the rat approaches the food well during presentations of the CS, and another is called sign-tracking (e.g., Hearst & Jenkins, 1974), where the rat interacts with the lever. It transpires that if a group of rats is given such training, then different rats will exhibit knowledge about the lever-food relationship in distinct ways: Some predominantly interact with the lever, others predominantly approach the food well, and the remainder exhibit more similar levels of the two behaviors (e.g., Iliescu, Hall, Wilkinson, Dwyer & Honey, 2018; Flagel, Akil & Robinson, 2009; Patitucci, Nelson, Dwyer & Honey, 2016; cf. Matzel et al., 2003). This variation in conditioned behavior across rats is remarkably stable from one day of training to the next, and is thus appropriately described as an individual difference. Most formal models of Pavlovian conditioning are not well placed to explain such qualitative differences in performance. They assume that the relationship between the strength of a CS-US association ( $V_{CS-US}$ ) and performance is monotonic (e.g., Rescorla & Wagner, 1972). This assumption does not allow a single acquired property (like associative strength,  $V$ ) to be manifest in quite different ways across a group of rats. We have recently described a theoretical model (HeiDI), which attempts to address this issue, alongside other challenging theoretical issues (Honey, Dwyer & Iliescu, 2020ab).

HeiDI uses trial-based learning rules to describe the formation of reciprocal associations between the representations of the CS and US; and a detailed description of its relationship with the model described by Rescorla and Wagner (1972; Wagner & Rescorla, 1972) can be found in Honey et al. (2020ab). However, the principal differences between HeiDI and the Rescorla-Wagner model are: HeiDI specifies rules governing the formation of the reciprocal (CS-US and US-CS) associations (Equations 1 and 2) and their combination (Equation 3) whereas the Rescorla-Wagner model did not; and HeiDI specifies how this combined associative strength is distributed into two components that affect CS-oriented behavior (e.g., sign-tracking) and US-oriented behavior (e.g., goal tracking; Equations 4 and 5), whereas the Rescorla-Wagner model assumed a simple ordinal relationship between the strength of the CS-US association and conditioned responding on all behavioral measures. In Equation 1, the maximum strength of the CS-US association is 1 in units of  $V$  (denoted  $c$ ), which is modulated by the value of the parameter  $\beta_{US}$  that is aligned to the perceived salience of the US (i.e.,  $c \cdot \beta_{US}$ ). In Equation 2, the maximum strength of the US-CS association is also 1 in units of  $V$  (denoted  $c$ ), which is modulated by the value of the parameter  $\alpha_{CS}$  that is aligned to the perceived salience of the CS (i.e.,  $c \cdot \alpha_{CS}$ ). Changes in the strength of the CS-US association ( $\Delta V_{CS-US}$ ) and US-CS association ( $\Delta V_{US-CS}$ ) are held to be determined by the value of the learning rate parameters for the CS ( $\alpha_{CS}$ ) and US ( $\beta_{US}$ ) multiplied by the respective pooled error terms:  $(c \cdot \beta_{US} - \Sigma V_{TOTAL US})$  and  $(c \cdot \alpha_{CS} - \Sigma V_{TOTAL CS})$ .  $\Sigma V_{TOTAL US}$  and  $\Sigma V_{TOTAL CS}$  represent the combined associative strengths of the stimuli that are present on a given trial with respect to the subscripted stimulus (i.e., US and CS, respectively). In both equations, the learning rate parameters ( $\alpha_{CS}$  and  $\beta_{US}$ ) are confined to the unit interval:  $0 \leq \alpha_{CS}, \beta_{US} \leq 1$ ; and both are dimensionless scalars roughly aligned to the (perceived) salience of the stimuli. The model assumes that the perceived salience of a given stimulus (CS and US) can vary between different animals, but is fixed for a given animal. This assumption provides grounds for individual differences in both the strength and nature of

conditioned responding once associative strength is coupled with suitable performance rules, which we shall come to shortly (Equations 4 and 5). Returning to Equations 1 and 2, learning ceases when the aggregated associative strengths (e.g.,  $\Sigma V_{\text{TOTAL-US}}$ ) equals the asymptote determined by the target of the association (e.g.,  $c \cdot \beta_{\text{US}}$ ).

$$\Delta V_{\text{CS-US}} = \alpha_{\text{CS}} (c \cdot \beta_{\text{US}} - \Sigma V_{\text{TOTAL-US}}) \quad (1)$$

$$\Delta V_{\text{US-CS}} = \beta_{\text{US}} (c \cdot \alpha_{\text{CS}} - \Sigma V_{\text{TOTAL-CS}}) \quad (2)$$

The strengths of the reciprocal associations are combined according to Equation 3, and the resulting  $V_{\text{COMB}}$  is distributed into CS-oriented responding ( $R_{\text{CS}}$ ; e.g., sign-tracking) and US-oriented responding ( $R_{\text{US}}$ ; e.g., goal-tracking) according to Equations 4 and 5, respectively. In Equation 3,  $V_{\text{COMB}}$  combines  $V_{\text{CS-US}}$  with the product of  $V_{\text{CS-US}}$  and  $V_{\text{US-CS}}$ ; the reciprocal of  $c$  is used to convert  $V_{\text{CS-US}}$  into a dimensionless scalar. The multiplicative term in Equation 3 means that the link directly activated by the presentation of the CS will constrain the impact of the indirectly activated reciprocal link. According to Equations 4 and 5, the distribution of  $V_{\text{COMB}}$  into  $R_{\text{CS}}$  and  $R_{\text{US}}$  is determined by the value of  $\alpha_{\text{CS}}$  relative to  $V_{\text{CS-US}}$ . That is, this distribution is determined by the perceived salience of the CS ( $\alpha_{\text{CS}}$ ) relative to its capacity to activate the US representation ( $V_{\text{CS-US}}$ ); with the absolute value being used to ensure that the proportions take values  $\leq 1$  even when  $V_{\text{CS-US}}$  takes a negative value. As before,  $|V_{\text{CS-US}}|$  is transformed into a dimensionless value by multiplying it by  $1/c$ , which means that  $R_{\text{CS}}$  and  $R_{\text{US}}$  are in units of  $V$ .

$$V_{\text{COMB}} = V_{\text{CS-US}} + \left( \frac{1}{c} \cdot V_{\text{CS-US}} \times V_{\text{US-CS}} \right) \quad (3)$$

$$R_{\text{CS}} = \frac{\alpha_{\text{CS}}}{\alpha_{\text{CS}} + \frac{1}{c} \cdot |V_{\text{CS-US}}|} V_{\text{COMB}} \quad (4)$$

$$R_{\text{US}} = \frac{\frac{1}{c} \cdot |V_{\text{CS-US}}|}{\alpha_{\text{CS}} + \frac{1}{c} \cdot |V_{\text{CS-US}}|} V_{\text{COMB}} \quad (5)$$

The analysis offered by HeiDI for individual variation in the form of conditioned responding rests on the assumption that there are individual differences in  $\alpha_{CS}$ , which directly affect the values returned by Equations 4 and 5, and differences in  $\beta_{US}$ , which indirectly affect the values returned by Equations 4 and 5 through their impact on  $V_{CS-US}$ . Because  $V_{CS-US}$  is influenced both by the asymptote that  $c \cdot \beta_{US}$  sets, and through the values returned by Equation 1, then  $R_{CS}$  and  $R_{US}$  can be said to be related to the perceived salience of the associatively activated US representation (determined by  $V_{CS-US}$ ). In any case, as  $\alpha_{CS}$  increases relative to  $V_{CS-US}$ ,  $R_{CS}$  increases relative to  $R_{US}$ , and as  $V_{CS-US}$  increases relative to  $\alpha_{CS}$ ,  $R_{US}$  increases relative to  $R_{CS}$ .

As already noted in passing, HeiDI is currently formulated as a trial-based model: The associative changes ( $\Delta V_{CS-US}$  and  $\Delta V_{US-CS}$ ) described by Equations 1 and 2 ignore both the temporal order of the components of the trial, and the fact that the components are extended in time (see also, Lesaint, Sigaud, Flagel, Robinson, & Khamassi, 2014). For this reason, the model does not capture variation in the tendency of a CS to elicit different forms of behavior ( $R_{CS}$  and  $R_{US}$ ) across its temporal extent. The limited evidence regarding such variation is inconsistent. In one study, there was a brief spike in goal-tracking within the first two seconds of lever onset, which was evident when the intertrial interval was 60s, but not when it was 120s; and sign-tracking increased progressively across the duration of the CS irrespective of the intertrial interval (Lee et al., 2018). These results are rather surprising: Previous reports have shown that goal-tracking increases across the duration of a CS (see Holland, 1977; Nasser, Chen, Fiscella & Calu, 2015; see for further discussion and analysis, Cinotti, Marchand, Roesch, Girard & Khamassi, 2019). Indeed, it is well established that as training progresses, conditioned responding becomes increasingly evident towards the end of a CS (an effect dubbed *inhibition of delay* by Pavlov, 1927; see pp. 61-62, Mackintosh, 1974). In the only other study of this kind, the small subset of rats that engaged in both goal-tracking and sign-tracking showed an increase in goal-tracking and a decrease in sign-tracking across the lever presentations (Derman,



Schneider, Juarez, & Delamater, 2018; see Figures 1b and 3c, noting the corrigendum confirming that the legend for Figure 1b was reversed for the right-hand panel). In the majority of the rats, who were sign-trackers, lever presses initially increased and then decreased across the presentation of the lever. While any increase in responding over the first few seconds of a CS is likely to reflect, at least in part, the time taken to move to the manipulandum (e.g., the lever), the decline in this response towards the end of the lever presentation matches that seen the subset of rats that engaged in some goal-tracking. The pattern of results observed in the small subset of the rats is of theoretical interest, as is the decline in conditioned lever presses that was observed in the majority of rats.

There are several interpretations of inhibition of delay. For example, some models attribute such temporal control of conditioned behaviors to a process of timing (e.g., Gallistel & Gibbon, 2000; see also Savastano & Miller, 1998), while others assume that a CS can be broken down into a succession of overlapping elements that gain and lose associative strength according to standard associative rules (e.g., Vogel, Brandon & Wagner 2003). But, these models do not address either how different forms of conditioned responding might change across a CS presentation, or how such responses might vary across individuals. In the same way that individual differences in sign-tracking and goal-tracking are not accommodated by models in which a single process (timing or association) maps onto conditioned performance, differential changes in these conditioned behaviors across the duration of a CS are also problematic: How could a single process (e.g., associative strength) affect two conditioned behaviors in different ways across the duration of the CS?

The aims of the current analyses were twofold: To assess how sign-tracking and goal-tracking vary across the duration of a CS (cf. Derman et al., 2018; Lee et al., 2018); and to determine whether a simple modification to HeiDI allows it to provide an account of such variation, which is routinely taken to reflect a process of timing (e.g., Gallistel & Gibbon, 2000;

see also Savastano & Miller, 1998) or the discrimination of successive parts of a CS (e.g., Vogel et al., 2003). We examined the empirical issue using archival data from Iliescu et al. (2018). The simple modification to HeiDI that we considered was to allow the perceived salience of the CS (i.e.,  $\alpha_{CS}$ ) to decline across its duration (e.g., according to an exponential decay function). This idea was briefly noted by Mackintosh (p. 62, 1974; see also, Pavlov, 1927, p. 104; Staddon, 2005; Staddon & Higga, 1999; Wagner, 1981), who argued that such a decay process might enable the early and later parts of a CS to be discriminated, and thereby provide a basis for inhibition of delay. However, within HeiDI, the effects of introducing a reduction in the perceived salience across a CS on learning and performance are more straightforward. First, such a reduction would result in poorer learning to a CS that is extended in time, to the extent that the  $\alpha_{CS}$  value at the point of US delivery is lower than at CS onset (in Equations 1 and 2). Second, the dynamic changes in  $\alpha_{CS}$  across successive epochs of the CS could affect performance through application of Equations 4 and 5. In fact, as we will confirm, this simple change results in a decline in sign-tracking and an increase in goal-tracking across a CS; with these changes interacting with the initial levels of  $\alpha_{CS}$  relative to  $\beta_{US}$ . It will be remembered that this was the pattern of results consistently observed in a small subset of the rats reported by Derman et al. (2018), but not that observed by Lee et al. (2018). We will present formal simulations of the pattern of results predicted by HeiDI once the re-analysis of our empirical results has been presented.

## Experiments 1 and 2

Our analysis of the distribution of sign-tracking and goal-tracking across the duration of a CS was based on an analysis of archival data from two groups of rats given similar training, and taken from Experiment 1 (female rats) and Experiment 2 (male rats) in Iliescu et al. (2018). Rats in Experiment 1 received 10-s presentations of one lever ( $L_1$ ; e.g., the lever positioned to the left of a food well) paired with the delivery of a single food pellet, and 10-s presentations of a second

lever (L<sub>2</sub>; e.g., the lever positioned to the right of a food well) that were not followed by food. For those in Experiment 2, both levers were paired with food. Automated recordings of the interactions of rats with the levers (i.e., sign-tracking) and the food-well (i.e., goal-tracking) were used to classify the rats as sign-trackers (ST) and goal-trackers (GT). Here, the rats were classified in the same way as in Iliescu et al. (2018). This classification was based on whether their behavior was predominantly oriented towards the lever (group ST) or food well (group GT). This bias is relatively continuous across our rats (see Figures 2 and 4, Iliescu et al., 2018). Dividing the rats into two relatively large groups on the basis of this bias has the advantage of including all of the rats in the analysis, but it inevitably means that the distinction between sign-trackers and goal-trackers is not all-or-none. Thus, both groups ST and GT will include rats that would be classified as ‘intermediates’ (e.g., Derman et al., 2018). Our primary interest was to assess the distribution of the two forms of behavior across the 10-s lever presentations in groups ST and GT. This assessment was conducted by dividing the 10-s presentations into successive 2.5-s epochs, balancing the need for temporal granularity with securing a representative sample of behavior.

## Method

*Subjects.* The procedures for Experiments 1 and 2 were described in Iliescu et al. (2018). Sixteen female Sprague Dawley rats (mean ad lib weight = 321g; range: 280-366g; supplied by Charles River, UK) were used in Experiment 1. They were naïve with respect to the apparatus and procedures, but had served in a behavioral task involving drinking different concentrations of sucrose. Sixteen naïve male (outbred) Lister Hooded rats (mean ad lib weight = 317g; range: 284-340g; these are a subset that received food pellets as in Experiment 1; supplied by Envigo, UK) were used in Experiment 2.<sup>1</sup> In both experiments, rats were housed in groups ranging from

---

<sup>1</sup> The rats given sucrose as the US were excluded in this analysis, because their levels of responding were lower, making it difficult to detect changes in responding across the duration of the CS.

two to four in standard cages and maintained on 12-hr/12-hr light/dark cycle (lights on at 7 a.m.). The rats were maintained at between 85% and 95% of their ad lib weights by giving them restricted access to food at the end of each day in their home cages, where they had continuous access to water. The research was conducted in accordance with Home Office regulations under the Animal (Scientific Procedures) Act 1986.

*Apparatus.* Eight identical conditioning boxes (30×24×21 cm: H×W×D; Med Associates, Georgia, VT) were each placed in a sound-attenuating shell that incorporated a ventilation fan, which maintained the background noise at 68 dB(A). The side walls of the boxes were constructed from aluminum; the front, back and ceiling were made from clear acrylic; and the floor was formed from 19 steel rods (4.8 mm diameter, 16 mm apart) placed above a stainless-steel tray. Food pellets (45 mg; supplied by MLab; Richmond, IN) were delivered to a recessed food well (aperture: 5.3×5.3 cm) in the center of the left wall at floor level. The food well was equipped with infrared detectors that allowed the presence of the rat in the well to be automatically recorded. When the detector was interrupted (e.g., when a rat's snout entered the food well) a single response was registered. Two retractable levers (4.5×1.8×0.2 cm) were located 3 cm to the left and right of the food well. They were positioned at a height of 4.6 cm and 1.5 cm from the edge of the wall. Depression of the lever by 4mm from its usual horizontal resting position was recorded as a response. MED-PC software was employed to control the insertion of levers, delivery of food pellets, and to record food well entries and lever presses.

*Procedure.* The procedures for Experiments 1 and 2 were described in Iliescu et al. (2018). Rats had two 24-min pre-training sessions in which food pellets were delivered on a variable-time (VT) 60-s schedule (range: 40–80s). On each of the next 12 days, they received a single training session, which occurred at the same time of day for a given rat. There were 40 trials in each session, 20 on which the left lever was inserted for 10s and then retracted, and 20 trials on which the right lever was inserted for 10s and was then withdrawn. In Experiment 1, for half of

the rats, the retraction of the left lever was immediately followed by the delivery of one food pellet and presentations of the right lever were not reinforced; and for the other half this arrangement was reversed. In Experiment 2, both levers were followed by the delivery of a food pellet. The order in which the two levers were presented was random with the constraint that there were no more than three presentations of the same lever in succession. The trials were delivered on a variable-time (VT) 60-s schedule (range: 40–80s). There were additional manipulations in Iliescu et al. (2018; e.g., a change in the reinforcement contingencies after the 12 training sessions), but these are not described further because we only analyzed the first 12 training sessions.

*Data analysis.* Successive training sessions were combined into 6, 2-day blocks (T1-T6). At the end of the training phase, the rats were split into sign-trackers (group ST,  $ns=8$ ) and goal-trackers (group GT,  $ns=8$ ), based on their tendency to engage with the lever and the food well. A bias score was calculated using the number of lever presses and food-well entries for the reinforced lever,  $L_1$ :  $(\text{Goal-tracking} - \text{Sign-tracking}) / (\text{Goal-tracking} + \text{Sign-tracking})$ . A median split was used to divide rats into those with higher scores (group GT) and those with lower scores (group ST). One key aim of the current re-analysis of the data is to directly compare goal- and sign-tracking tendencies. However, using “raw” scores for both magazine entries and lever presses for this purpose is potentially problematic, because it requires the assumption that each lever press or magazine entry is equivalent in terms of reflecting sign- or goal-tracking strength. This assumption is questionable because there is little reason to think the mapping between associative strength and responses is the same for all response types (indeed, the full HeiDI model specifies separate equations for transforming  $R_C$ s and  $R_U$ s into observable responses, see Honey et al., 2020ab) and the raw results indicate that overall there are higher levels of food well entries than lever presses. Thus, in order to place lever-press and food-well

entry on the same scale, they were z-transformed prior to analysis<sup>2</sup>. z-scores were calculated separately for each response (lever-press and food-well entries). For lever presses this was equal to Value (a response level on a trial for a given rat) – Mean value of lever presses (across the 12 days and all rats), all divided by the standard deviation for lever presses (across all 12 days and rats). Similarly, for food-well entries this was equal Value (a response level on a trial for a given rat) – Mean value of food-well entries (across the 12 days and rats), all divided by the standard deviation for food-well entries (across all 12 days and rats).

An R script was used to clean the data provided by the MedPC. The script extracted the responses (lever presses and food-well entries) only during the CS period. The script for the data cleaning is available on the OSF ([https://osf.io/ka9w3/?view\\_only=753ea3b06ecd4ccf89550035bb60a6dc](https://osf.io/ka9w3/?view_only=753ea3b06ecd4ccf89550035bb60a6dc)). The 10-s CS period was divided into four 2.5-s intervals: epoch 1, epoch 2, epoch 3 and epoch 4 and averaged across the experimental day. For Experiment 1 this was calculated based on the response on the reinforced lever; and for Experiment 2 it was calculated as an average across both reinforced levers.

For manipulating the data, data cleaning, statistical analysis and data visualization, we used open source software R (RStudio, 2015). For data cleaning and manipulation, we used the “tidyr” (Wickham & Henry, 2018) and “dplyr” R packages (Wickham, Henry, & Muller, 2019). For hypothesis testing, we used the ‘ez’ R package (Lawrence, 2016). For data visualization and representation, we used the “ggplot2” (Wickham et al., 2018) R package, “cowplot” and “gridExtra” R packages (Auguie & Anotonov, 2017). The analyses were automatically reported in text in APA format style using the “apa” R package (Gromer, 2019). For analysis of variance

---

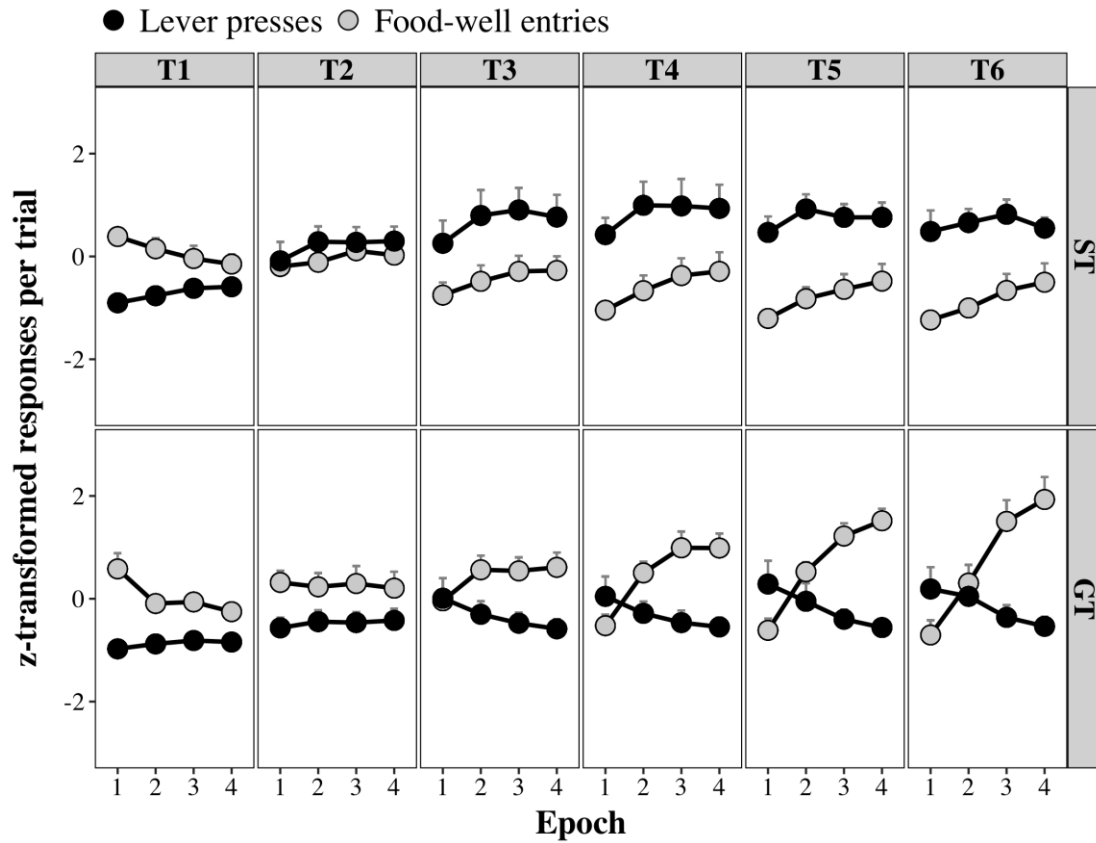
<sup>2</sup> Although only the transformed data are presented here, they key inferential results are the same for the raw and transformed data. The analysis of the untransformed data can be found in the supplementary materials.

(ANOVA) when the sphericity assumption was violated, the analysis was reported with Greenhouse-Geisser correction.

## Results

The distribution of sign-tracking and goal-tracking across the four successive (2.5-s) epochs of the 10-s reinforced lever presentations was assessed over the 6 blocks of training. These results are summarized in Figures 1 and 2. The six upper panels of Figure 1 (Experiment 1) and Figure 2 (Experiment 2) depict how the distribution of the two forms of response change across training in Group ST. The lower panels show the corresponding scores in Group GT. Lever press responses are indicative of sign-tracking behavior, while food-well entries reflect goal-tracking behavior. The results were similar in Experiments 1 and 2, with the exception that the levels of sign-tracking in group ST were lower in Experiment 1 (when one lever was reinforced and the other nonreinforced) than in Experiment 2 (when both levers were reinforced). This difference might reflect differences in the tendency to remain in the vicinity of the response panel as a consequence of the differing reinforcement contingencies. Inspection of both figures reveals that by the end of training, both ST and GT groups displayed a pattern of responding whereby food-well entries (i.e. goal-tracking behavior) increased relative to lever press responses (i.e. sign-tracking behavior) across the period of the lever CS presentation. However, by definition, ST groups displayed generally higher levels of lever press responses, and fewer food-well entries, than the GT groups, and so the expression of this general trend differed between them. Namely, as training proceeded in the ST groups, lever press responding was more evident than food-well entries early in the stimulus period, and that this difference was greatly reduced later in the stimulus primarily due to increases in food-well entries across the duration of the lever CS presentation (and indeed, the difference between lever press and food-well entries was absent or even reversed in Experiment 2). In contrast in the GT groups, as training progressed while lever press and food-well entries were similar early in the stimulus

period (with lever press numerically higher than food-well entries in the first epoch during the latter blocks), goal-tracking dominated sign-tracking during the latter part of the stimulus: with the levels of sign-tracking decreasing and the levels of goal-tracking increased.



*Figure 1.* Experiment 1. Mean (+SEM) z-transformed lever presses (black symbols) and food-well entries (grey symbols) during reinforced lever presentations in groups ST (upper panels) and GT (lower panels). The results from the 10-s lever presentations are broken down into 4 successive 2.5-s epochs for each of the 6 blocks of training (T1-T6).



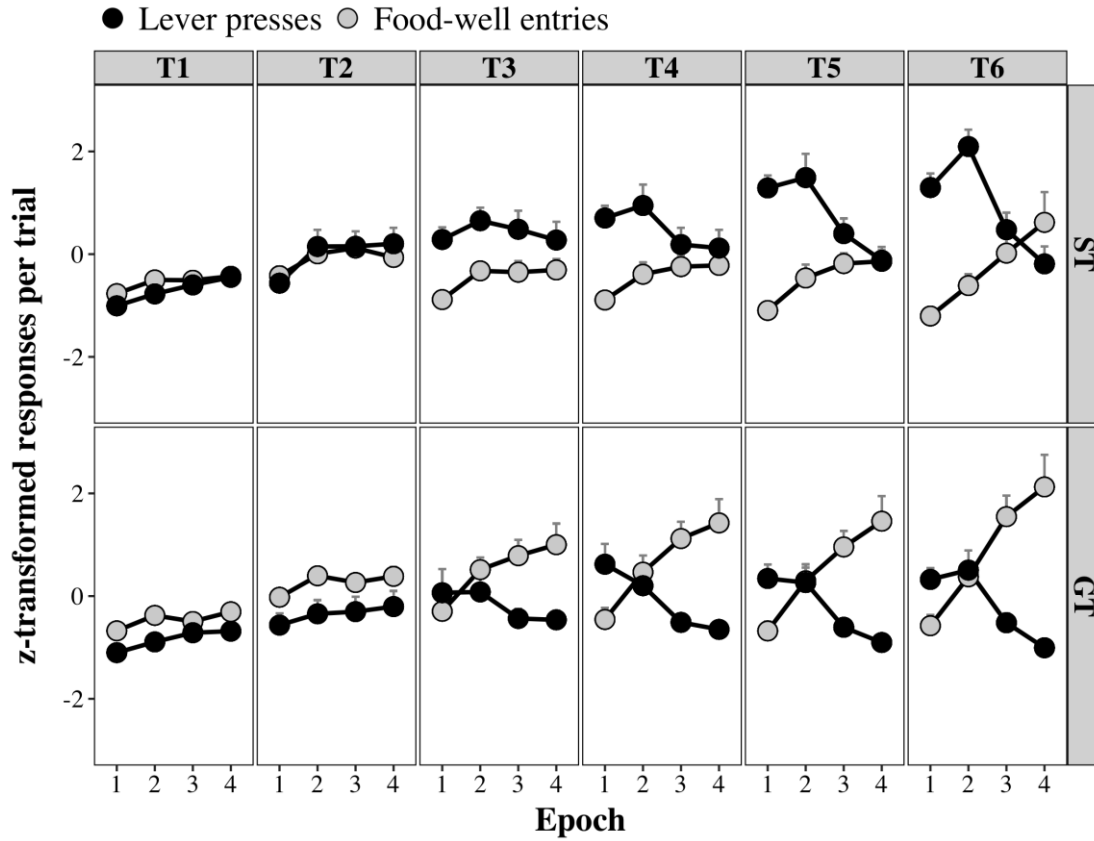


Figure 2. Experiment 2. Mean (+SEM) z-transformed lever presses (black symbols) and food-well entries (grey symbols) during reinforced lever presentations in groups ST (upper panels) and GT (lower panels). The results from the 10-s lever presentations are broken down into 4 successive 2.5-s epochs for each of the 6 blocks of training (T1-T6).

This description of the results was confirmed through mixed ANOVA analyses with training block (T1-T6), epoch (1-4), and response type (lever press or food-well entry) as within-subjects factors, and classification (group ST or GT) as the between-subjects factor. For Experiment 1, the fact that the distribution of lever presses and food-well entries changed across the duration of the lever presentation was reflected in an interaction between response type and epoch,  $F(1.135, 15.883) = 5.57$ ,  $p < .001$ ,  $\eta_p^2 = .28$ ,  $MSE = 0.96$ ; a pattern that evolved across training blocks, as indicated by a block by response type by epoch interaction,  $F(15, 210) = 9.03$ ,  $p < .001$ ,  $\eta_p^2 = .39$ ,  $MSE = 0.22$ . In turn, the change in the distribution of responses across the period of lever presentation differed between groups ST and GT as indicated by a classification

by response type by epoch interaction,  $F(1.13, 15.85) = 4.91$ ,  $p = .038$ ,  $\eta_p^2 = .26$ ,  $MSE = 0.96$ ; which also evolved across training blocks, as indicated by a classification by block by response type by epoch interaction,  $F(15, 210) = 3.24$ ,  $p < .001$ ,  $\eta_p^2 = .19$ ,  $MSE = 0.22$ .

The remainder of the ANOVA revealed main effects of block,  $F(2.84, 39.74) = 8.80$ ,  $p < .001$ ,  $\eta_p^2 = .39$ ,  $MSE = 6.23$  and epoch,  $F(1.26, 17.59) = 14.54$ ,  $p < .001$ ,  $\eta_p^2 = .51$ ,  $MSE = 0.33$ , but not of classification,  $F(1, 14) = 0.07$ ,  $p = .792$ ,  $\eta_p^2 < .01$ ,  $MSE = 6.23$ , or response type,  $F(1, 14) = 0$ ,  $p > .999$ ,  $\eta_p^2 < .01$ ,  $MSE = 8.8$ . There were two-way interactions between block and classification,  $F(2.84, 39.74) = 3.53$ ,  $p = .025$ ,  $\eta_p^2 = .20$ ,  $MSE = 0.5$ , response type and classification,  $F(1, 14) = 15.04$ ,  $p = .002$ ,  $\eta_p^2 = .52$ ,  $MSE = 8.8$  (reflecting the overall bias to lever press or food-well entry in ST and GT groups respectively), block and response type,  $F(2.52, 35.27) = 7.90$ ,  $p < .001$ ,  $\eta_p^2 = .36$ ,  $MSE = 0.92$ , block and epoch,  $F(5.46, 76.45) = 7.47$ ,  $p < .001$ ,  $\eta_p^2 = .35$ ,  $MSE = 0.1$ , but not between epoch and classification,  $F(1.26, 17.59) = 0.32$ ,  $p = .627$ ,  $\eta_p^2 = .02$ ,  $MSE = 0.33$ . There was also a significant 3-way interaction between block, response type and classification,  $F(2.52, 35.27) = 7.47$ ,  $p < .001$ ,  $\eta_p^2 = .35$ ,  $MSE = 0.92$ , but no block by epoch by classification interaction,  $F(5.46, 76.45) = 2.15$ ,  $p = .062$ ,  $\eta_p^2 = .13$ ,  $MSE = 0.1$ .

The most theoretically important result here is the fact that food-well entries (i.e. goal-tracking behavior) increased relative to lever presses (i.e. sign-tracking behavior) across the duration of the lever CS as training progressed. However, the 4-way interaction indicates that the distribution of responses across the period of lever presentation differed between ST and GT groups, and this difference evolved across blocks. Despite this, the general pattern of changes appears to be present in both ST and GT groups. This interpretation was supported by simple main effect analyses (performed as within-subjects  $t$ -tests). For block 1, food-well entries were significantly higher than lever presses in all epochs for group GT (smallest  $t(7) = 4.43$ ,  $p = .003$ ,  $d = 1.57$ ), while for group ST food-well entries were significantly higher than lever presses in the

first two epochs (smallest  $t(7) = 3.59$ ,  $p = .009$ ,  $d = 1.27$ ), but this difference was not significant in the last two epochs (largest  $t(7) = 2.08$ ,  $p = .076$ ,  $d = 0.74$ ). In contrast, the same analysis for block 6 revealed that for group ST, lever pressing was significantly higher than food-well entries in the first three epochs (smallest  $t(7) = -2.92$ ,  $p = .022$ ,  $d = -1.03$ ), but not in the fourth epoch ( $t(7) = -2.14$ ,  $p = .070$ ,  $d = -0.76$ ); whereas for group GT there was no significant difference between lever presses and food-well entries in the first or second epochs (largest  $t(7) = -1.38$ ,  $p = .209$ ,  $d = -0.49$ ), while food-well entries were significantly higher than lever presses in the third and fourth epochs (smallest  $t(7) = 4.57$ ,  $p = .003$ ,  $d = 1.62$ ).<sup>3</sup> That is, by the end of training, both ST and GT groups displayed an increase in goal-tracking relative to sign-tracking across the duration of the lever that was not present at the start of training. For the ST group this was displayed as sign-tracking dominating goal-tracking early but not late in the stimulus period, while for the GT group it was displayed as goal-tracking dominating sign-tracking late but not early in the stimulus period.

A parallel ANOVA was conducted on the results from Experiment 2. Again, the fact that the distribution of lever-press and food-well entries changed across the duration of the lever was reflected in an interaction between response type and epoch,  $F(1.63, 22.88) = 22.19$ ,  $p < .001$ ,  $\eta_p^2 = .61$ ,  $MSE = 1.18$ ; a pattern that evolved across training blocks, as indicated by a block by

---

<sup>3</sup>Further statistical analyses conducted on Blocks 2-5 revealed the following patterns of statistical significance. Block 2: In group ST, the levels of the two responses did not differ during any epoch; whereas in group GT, there were more food well entries than lever presses during epoch 1 but not in other epochs. Blocks 3 and 4: In group ST, the levels of the two responses in block 3 did not differ during any epoch, whereas in block 4 there were more lever presses than food well entries during epochs 1 and 2, but no differences between the two responses during epochs 3 and 4. In group GT, during both blocks there was no difference between the two types of response during epoch 1 (or in epoch 2 of Block 3), but on the remaining epochs there were more food well entries than lever presses. Block 5: In group ST, there were more lever presses than food well entries during epochs 1-3, but no difference in the two responses during epoch 4; whereas in group GT the levels of the two responses did not differ during epochs 1 and 2, but differed thereafter.

response type by epoch interaction,  $F(3.28, 45.94) = 17.82, p < .001, \eta_p^2 = .56, \text{MSE} = 0.26$ . In Experiment 2, there was no classification by response type by epoch interaction,  $F(1.63, 22.88) = 1.66, p = .214, \eta_p^2 = .11, \text{MSE} = 1.18$ ; and no four-way interaction,  $F(3.28, 45.94) = 0.77, p = .528, \eta_p^2 = .05, \text{MSE} = 0.26$ .

The remainder of the ANOVA revealed main effects of block,  $F(2.58, 36.13) = 34.20, p < .001, \eta_p^2 = .71, \text{MSE} = 0.43$ , and epoch,  $F(1.83, 25.66) = 16.05, p < .001, \eta_p^2 = .53, \text{MSE} = 0.39$ , but not of classification,  $F(1, 14) = 0.20, p = .665, \eta_p^2 = .01, \text{MSE} = 5.17$ , or response type,  $F(1, 14) = 0, p > .999, \eta_p^2 < .01, \text{MSE} = 6.6$ . There were two-way interactions between classification and response type,  $F(1, 14) = 14.16, p = .002, \eta_p^2 = .50, \text{MSE} = 6.6$  (reflecting the overall bias to lever press or food-well entry in ST and GT groups respectively), block and response type,  $F(5, 70) = 2.72, p = .027, \eta_p^2 = .16, \text{MSE} = 0.55$ , block and epoch,  $F(4.91, 68.70) = 2.65, p = .031, \eta_p^2 = .16, \text{MSE} = 0.1$ , but not between epoch and classification,  $F(1.83, 25.66) = 2.05, p = .152, \eta_p^2 = .13, \text{MSE} = 0.39$ , or between block and classification,  $F(2.58, 36.13) = 0.97, p = .408, \eta_p^2 = .06, \text{MSE} = 0.43$ . There were, however, significant 3-way interactions between block, response type and classification,  $F(5, 70) = 8.78, p < .001, \eta_p^2 = .39, \text{MSE} = 0.55$ , and between block, epoch and classification,  $F(4.91, 68.70) = 2.40, p = .047, \eta_p^2 = .15, \text{MSE} = 0.1$ .

To understand the nature the 3-way interactions just reported for the results of Experiment 2, we conducted the same simple main effect analyses that were conducted for Experiment 1. During block 1, for group GT food-well entries were significantly higher than lever presses in epochs 1, 2, and 4 (smallest  $t(7) = 2.77, p = .028, d = 0.98$ ) and no difference in epoch 3 ( $t(7) = 2.04, p = .080, d = 0.72$ ); whereas for group ST, food-well entries did not differ significantly from lever presses in any epoch (largest  $t(7) = 1.59, p = .156, d = 0.56$ ). In contrast, the same analysis for block 6 revealed that for group ST lever presses were significantly higher than food-well entries in the first two epochs (smallest  $t(7) = -5.27, p = .001, d = -1.86$ ), but not in the third or fourth epochs (largest  $t(7) = 0.99, p = .353, d = 0.35$ ); whereas for group GT, lever presses

were higher than food-well entries in the first epoch ( $t(7) = -2.64$ ,  $p = .033$ ,  $d = -0.93$ ), there was no significant difference in the second epoch ( $t(7) = -0.23$ ,  $p = .823$ ,  $d = -0.08$ ), and there were more food-well entries than lever press rates in the third and fourth epochs (smallest  $t(7) = 4.79$ ,  $p = .002$ ,  $d = 1.69$ ).<sup>4</sup>

The results from Experiments 1 and 2 show that the distribution of sign-tracking and goal-tracking changes across a CS, and does so in a way that differs depending on whether a groups of rats has a propensity to sign-track (i.e., group ST) or to goal-track (i.e., group GT). At the start of a CS, rats in group ST were more likely to sign-track than to goal-track and this pattern became less evident by the end of the CS. Whereas for rats in group GT, there was little difference in their tendency to engage in sign-tracking or goal-tracking at the start of the CS, and their tendency to goal-track rather than sign-track emerged over the duration of the CS. One simple formal analysis for this intriguing pattern of results is presented in the immediately following discussion.

### General Discussion

Individual differences in Pavlovian conditioning have been ignored by formal models of associative learning (e.g., Mackintosh, 1975; Pearce & Hall, 1980; Rescorla & Wagner, 1972; Stout & Miller, 2007; Wagner, 1981): Treated as little more than a nuisance variable, their nature

---

<sup>4</sup>Further statistical analyses conducted on Blocks 2-5 revealed the following patterns of statistical significance. Block 2: In group ST, the levels of the two responses did not differ during any epoch; whereas in group GT, there were more food-well entries than lever presses during epoch 2 but not in other epochs. Blocks 3 and 4: In group ST, there were more lever presses than food-well entries during epoch 1 (and epoch 2 of block 3), but no differences between the two responses elsewhere. In group GT, during both blocks there was no difference between the two types of response during epochs 1 or 2, but on the remaining epochs there were more food-well entries than lever presses. Block 5: In group ST, there were more lever presses than food-well entries during epochs 1 or 2, but no difference in the two responses during epoch 3 or 4; whereas in group GT the levels lever press were higher than food-well entries in epoch 1, the levels of the two responses did not differ during epoch 2, and food-well entries were higher than lever press in epochs 3 and 4.

has been left unconsidered. However, the fact that different measures of conditioned performance (e.g., sign-tracking and goal-tracking) can provide the basis for drawing opposing inferences about  $V$  is something that any theory of associative learning should be concerned with: Is  $V$  greater in rats designated as sign-trackers or those designated as goal-trackers? This question is clearly not sensible (cf. Akins, Domjan & Gutierrez, 1994). One needs a model in which the relationship between  $V$  and different forms of conditioned behavior is both specified and provides an analysis of the basis for individual differences in those behaviors.

We have developed a model of Pavlovian conditioning, HeiDI, which addresses how associative strength is distributed between different forms of conditioned responding (CS-oriented or US-oriented; Honey et al., 2020ab). The critical assumption, which enables HeiDI to provide an analysis of individual differences in the form of conditioned responding, is that the perceived salience of the CS (i.e.,  $\alpha_{CS}$ ) and US ( $\beta_{US}$ ) differs between animals. Once these differences are coupled with appropriate learning and performance rules (Equations 1-5), a coherent analysis can be developed for variation in the strength and form of conditioned behavior across different animals. Briefly, if  $\alpha_{CS}$  is high relative to  $\beta_{US}$  (and  $V_{CS-US}$ ) then CS-oriented behavior dominates US-oriented behavior, and if  $\beta_{US}$  (and  $V_{CS-US}$ ) is high relative to  $\alpha_{CS}$  then US-oriented behavior dominates CS-oriented behavior; and when the two parameters are relatively similar then the levels of CS-oriented and US-oriented behavior reflect this similarity. However, because HeiDI is a trial-based model it does not capture the temporal control of behavior across the duration of a CS: if one measures conditioned responding across a CS one finds that conditioned responding increases over its duration. This effect is known as inhibition of delay (Pavlov, 1927). Here, we analyzed how sign-tracking and goal-tracking change across the duration of a CS.

The pattern of results was clear: Rats designated as sign-trackers were more likely to exhibit their bias to sign-track at the start of the CS, while rats designated as goal-trackers were

more likely to exhibit their bias to goal-track toward the end of the CS. At no point during the 6 blocks of training did the pattern of results observed in the groups designated as sign-trackers and goal-trackers match that reported in Lee et al. (2018). However, they did match closely those reported by Derman et al. (2018), especially in the small subset of rats that engaged in both sign-tracking and goal-tracking. The basis for these differing patterns of results is not clear, and any analysis would be necessarily speculative. For example, there is certainly between-strain variability in the extent to which sign-tracking and goal-tracking are evident (for a review, see Fitzpatrick et al., 2013); and unlike our rats, those from the Lee et al. (2018) study showed more sign-tracking than goal-tracking. However, the majority of rats from Derman et al. (2018) showed more sign-tracking than goal-tracking and yet the pattern of results that they observed was similar to that observed here. Similarly, while Experiments 1 and 2 used 2.5-second epochs, both Lee et al. (2018) and Derman et al. (2018) used 1-second epochs, and yet the results of Experiments 1 and 2 were similar to the latter, and both differed from the former. Perhaps the most obvious procedural difference between Lee et al. (2018), on the one hand, and the studies reported here and by Derman et al (2018), on the other hand, is that only in Lee et al. (2018) did rats wear headpieces necessary to assess dopamine using fast scan cyclic voltammetry. This addition might have interacted with their willingness to place their heads into a small aperture during the CS. There are, after all, other demonstrations that standard inhibition of delay effects can be observed using food-well entries (e.g., Delamater & Holland, 2008; Kirkpatrick & Church, 2004).

Before demonstrating how a simple adaptation of HeiDI enabled it to simulate the patterns of sign- and goal-tracking observed in sign-trackers and goal-trackers in Experiments 1 and 2, we will first briefly consider alternative accounts. One seductive account for the results of Experiments 1 and 2 rests on the idea that the two responses (lever pressing and food-well entries) simply compete with one another at the level of behavioral output: The rats cannot

readily interact with the lever and food well at the same time (e.g., Dwyer, Starns & Honey, 2009). A moment's reflection, however, reveals the inadequacy of such an analysis: Why does sign-tracking compete more effectively with goal-tracking at the onset of the presentation of the lever, but less effectively at the end? One could make some additional assumptions. For example, one might assume that the early parts of the lever presentation have lower associative strength than the later parts (e.g., Vogel et al., 2003), and that lower associative strength gives rise to sign-tracking and higher levels of associative strength to goal-tracking. However, one would still need to specify why it is that lower levels of associative strength are evident as sign-tracking and higher levels as goal-tracking. Indeed one would also need to explain the fact that over the course of training, when associative strength is increasing, sign-tracking increases. This form of explanation for the results of Experiments 1 and 2 is superficially seductive: It appears simple but it is in fact difficult to develop in a way that is consistent with the results (for further discussion of this issue, see Honey, Dwyer & Iliescu, 2020c).

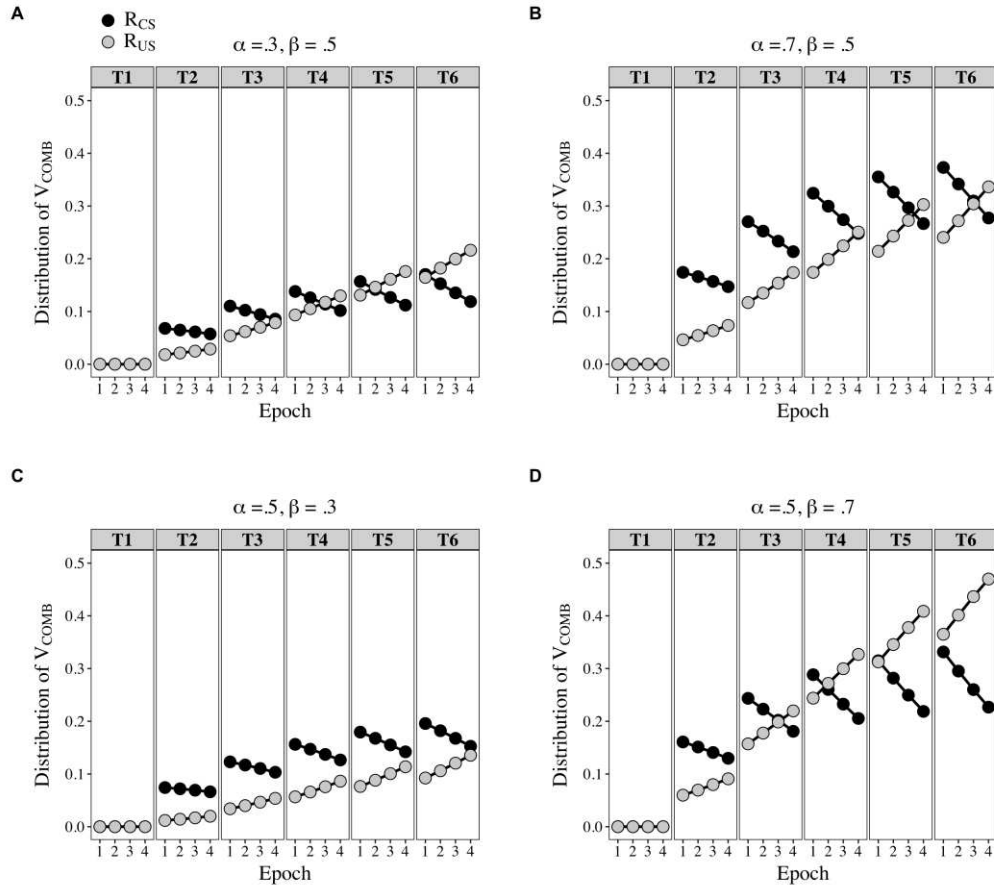
Another potential account aligns different behaviors (sign-tracking and goal-tracking) to different learnt content. For example, Timberlake (1993) has advocated a behavior systems approach in which "Learning is assumed to occur at many points in a behavior system, in varied forms, and as a function of conditions ranging from simple stimulus exposure to explicit response contingencies." (p. 118-119); and proceeded to argue that "it would not be surprising to find "causality detectors" (one or more mechanisms for indexing contingencies between stimuli) in most systems and species, although it would be surprising if they were identical in operation." (p. 122). This approach is clearly well equipped to deal with multiple forms of learnt behaviors originating from different parts of a behavior system. Similarly, one might align sign-tracking to preparatory conditioning, insofar as it is more marked at the start of a CS than the end, and goal-tracking to consummatory conditioning (Konorski, 1967). However, neither the behavior systems approach nor the appeal to the distinction between consummatory and preparatory conditioning



provides an obvious analysis for individual differences in sign-tracking and goal-tracking. We now return to a model that does provide such an analysis: HeiDI.

In the introduction we noted that one way in which HeiDI could be modified is to let the perceived salience of the CS ( $\alpha_{CS}$ ) decline across its duration, and then to re-set at the start of the next trial (p. 62, Mackintosh, 1974; see also, Pavlov, 1927, p. 104; Staddon, 2005; Staddon & Higga, 1999; Wagner, 1981). For the sake of simplicity and for the time being, we will assume that re-setting is complete and the degree of salience decline consistent, while acknowledging it might interact with details of the conditioning procedure (e.g., the intertrial interval or the length of the CS). We will now demonstrate how simply allowing the perceived salience of a CS to decline across its duration enables HeiDI to capture – to a first approximation – the results of Experiments 1 and 2. In doing so, we show how a phenomenon that seems to reflect a process of timing (i.e., inhibition of delay) can be explained more economically. The upper panels of Figure 3 show the results of simulations in which the initial  $\alpha_{CS}$  value (i.e., at the start of the CS) was set at either 0.30 (left panel) or 0.70 (right panel), while  $\beta_{US}$  was fixed at 0.50; and the lower panels show the results of simulations in which  $\alpha_{CS}$  was set at 0.50 (i.e., at the start of the CS) and  $\beta_{US}$  was set at either 0.30 (left panel) or 0.70 (right panel). The  $\alpha_{CS}$  value in the fourth epoch was used to calculate the change in associative strength in Equations 1 and 2. If we ignore, for the time being, variations in  $R_{CS}$  and  $R_{US}$  across the epochs within a trial, then the patterns of changes in  $R_{CS}$  and  $R_{US}$  matches those observed when there is no  $\alpha_{CS}$  decay function. Thus, if  $\alpha_{CS}$  is set at a higher value than  $\beta_{US}$  (panels B and C), then the values returned for  $R_{CS}$  tend to be larger than those returned for  $R_{US}$ ; but if  $\beta_{US}$  is set at a higher value than  $\alpha_{CS}$  (panels A and D), then the values returned for  $R_{US}$  tend to be larger than those returned for  $R_{CS}$  (cf. Honey et al., 2020a). One interesting (supplementary) observation is that in each case while the values returned for  $R_{CS}$  increase according to a negatively accelerated function, those returned for  $R_{US}$  increase according to a sigmoidal function. This observation reflects the fact while the value of

$\alpha_{CS}$  is substituted into Equations 4 and 5, the influence of varying  $\beta_{US}$  is through  $|V_{CS-US}|$ , which increases across training trials. This has the consequence that the distribution of  $V_{COMB}$ , which itself rises according to a negatively accelerating function reflecting  $V_{CS-US}$  and  $V_{US-CS}$ , is biased towards  $R_{CS}$  at the start of training.



*Figure 3. Simulations of performance across conditioning trials.* The  $V_{COMB}$  outputs used in Equations 4 and 5 to calculate  $R_{CS}$  (black symbols) and  $R_{US}$  (grey symbols) were generated using Equations 1-3. In panels A and B, at the start of a CS presentation  $\alpha_{CS}$  was either .3 (A) or .7 (B) and  $\beta_{US}$  was fixed at .5; and in panels C and D, at the start of the CS presentation  $\alpha_{CS}$  was fixed at .5 and  $\beta_{US}$  was either .3 (C) or .7 (D). Across the 4 epochs of a trial, the value of  $\alpha_{CS}$  was subject to exponential decay:  $\alpha_{CS}(1 - 0.10)^2$ . The terminal values of  $\alpha_{CS}$  were used in Equations 1 and 2, while the (within-trial) decaying values were used in Equations 4 and 5.

Of most interest here, however, is the way in which allowing the initial  $\alpha_{CS}$  values to be subject to a process of adaptation or decay affects the pattern of performance across the duration

of a CS. In the simulations shown in Figure 3 the exponential decay function was:  $\alpha_{CS}(1 - 0.10)^2$ . When the starting values of  $\alpha_{CS}$  were less than those of  $\beta_{US}$  (panels A and D), then as training progressed the tendency for the values returned for  $R_{US}$  to exceed those for  $R_{CS}$  increases over the duration of the CS. In contrast, when the starting values of  $\beta_{US}$  were less than  $\alpha_{CS}$  (panels B and C), then as training progressed the tendency for the values returned for  $R_{CS}$  to exceed those for  $R_{US}$  decreases over the duration of the CS. To summarize: While a bias towards  $R_{CS}$  (reflected in sign-tracking) is predicted to decrease across the duration of a CS, a bias towards  $R_{US}$  (reflected in goal-tracking) is predicted to become more evident over the duration of a CS. This is the pattern of results that we observed towards the end of training in Experiments 1 and 2 (see Figures 1 and 2; see Derman et al., 2018; but see Lee et al., 2018).<sup>5,6</sup>

Our new results show that individual variation in sign-tracking and goal-tracking change in orderly ways across the duration of the CS: A sign-tracking bias becomes less apparent and goal-tracking bias becomes more evident across a CS. HeiDI provides an account of these biases within a general process associative framework through assuming that the perceived salience of the CS and US affect performance as well as learning. If one adopts the plausible assumption that the perceived salience of a stimulus decreases across its duration, then HeiDI also provides a simple analysis for how these biases change across the presentation of the CS.

---

<sup>5</sup>The fact that all rats tend to begin training with a bias to enter the food well, which probably reflects the influence of their pre-training to retrieve food pellets from the food well, makes it difficult to evaluate whether the simulations for the early part of training are accurate or not.

<sup>6</sup>Further simulations revealed that increasing the steepness of the exponential function reduced  $V_{COMB}$ , because of its influence on the value of  $\alpha_{CS}$  used in Equations 1 and 2, and increased the bias towards  $R_{US}$  in general and across CS duration.

#### Author note

The original empirical research was conducted when A.F.I. was supported by a School of Psychology PhD studentship, and supervised by D.M.D. and R.C.H. It was also supported by a Strategic Award from the Wellcome Trust (100202/Z/12/Z), on which Jeremy Hall and Lawrence S. Wilkinson were PIs. The development of HeiDI and the current re-analysis of the archival data was supported by a BBSRC (UK; BB/T004339/1) grant awarded to R.C.H. and D.M.D. A.F.I. conducted the original research, A.F.I., D.M.D., and R.C.H. conceived the study, performed the statistical analyses, conducted the simulations, and wrote the paper.

## References

- Akins, C.K., Domjan, M., and Gutierrez, G. (1994). Topography of sexually conditioned behavior in male Japanese quail (*Coturnix japonica*) depends on the CS-US interval. *Journal of Experimental Psychology: Animal Behavior Processes*, 20, 199-209.
- Auguie, B., & Anotonov, A. (2017). package `gridExtra`: Miscellaneous Functions for “Grid” Graphics. Retrieved from <https://cran.r-project.org/web/packages/gridExtra/index.html>
- Boakes, R.A. (1977). Performance on learning to associate a stimulus with positive reinforcement. In H. Davis, & H.M.B. Hurwitz (Eds). *Operant-Pavlovian interactions* (pp. 67-97). NJ: Lawrence Erlbaum Associates.
- Cinotti, F., Marchand, A.R., Roesch, M.R., Girard, B., & Khamassi, M. (2019). Impacts of inter-trial interval duration on a computational model of sign-tracking vs. goal-tracking behaviour. *Psychopharmacology*, 236, 2373-2388.
- Delamater, A. R., & Holland, P. C. (2008). The Influence of CS-US Interval on Several Different Indices of Learning in Appetitive Conditioning. *Journal of Experimental Psychology: Animal Behavior Processes*, 34, 202-222.
- Derman, R.C., Schneider, K., Juarez, S, & Delamater, A.R. (2018). Sign-tracking is an expectancy-mediated behavior that relies on prediction error mechanisms. *Learning & Memory*, 550-563.
- Dwyer, D.M., Starns, J., & Honey, R.C. (2009). "Causal reasoning" in rats: A re-appraisal. *Journal of Experimental Psychology: Animal Behavior Processes*, 35, 578-586.
- Fitzpatrick, C.J., & Church, R.M. (2004). Temporal learning in random control procedures. *Journal of Experimental Psychology: Animal Behavior Processes*, 30, 213-228.

- Fitzpatrick, C.J., Gopalakrishnan, S., Cogan, E.S., Yager, L.M., Meyer, P.J., Lovic, V., Saunders, B.T., Parker, C.C., Gonzales, N.M., Aryee, E., Flagel, S.B., Palmer, A.A., Robinson, T.E., & Morrow, J.D. (2013). Variation in the form of Pavlovian conditioned approach behavior among outbred male Sprague-Dawley rats from different vendors and colonies: Sign-tracking vs. goal-tracking. *PLoS ONE*, [10.1371/journal.pone.0075042](https://doi.org/10.1371/journal.pone.0075042).
- Flagel, S.B., Akil, H., & Robinson, T.E. (2009). Individual differences in the attribution of incentive salience to reward-related cues: Implications for addiction. *Neuropharmacology*, *56*, 139-148.
- Gallistel, C.R., & Gibbon, J. (2000) Time, rate and conditioning. *Psychological Review*, *107*, 289-344.
- Gromer, D. (2019). package ‘apa’: Format Outputs of Statistical Tests According to APA Guidelines. Retrieved from <https://cran.r-project.org/web/packages/apa/index.html>
- Hearst, E., & Jenkins, H. M. (1974). *Sign tracking: The stimulus-reinforcer relation and directed action*. Austin: TX: Psychonomic Society.
- Holland, P.C. (1977). Conditioned stimulus as a determinant of the form of the Pavlovian conditioned response. *Journal of Experimental Psychology: Animal Behavior Processes*, *3*, 77–104.
- Honey, R.C., Dwyer, D.M., & Iliescu, A.F. (2020a). HeiDI: A model for Pavlovian learning and performance with reciprocal associations. *Psychological Review* (in press).
- Honey, R.C., Dwyer, D.M., & Iliescu, A.F. (2020b). Elaboration of a model of Pavlovian learning and performance: HeiDI. *Journal of Experimental Psychology: Animal Learning and Cognition* (in press).
- Honey, R.C., Dwyer, D.M., & Iliescu, A.F. (2020c). Individual variation in vigor and form of Pavlovian conditioned responses. *Learning & Motivation* (in press).

- Iliescu, A.F., Hall, J., Wilkinson, L., Dwyer, D.M., & Honey, R.C. (2018). The nature of phenotypic variation in Pavlovian conditioning. *Journal of Experimental Psychology: Animal Learning and Cognition*, 44, 358-369.
- Jeffreys, H. (1961). *Theory of Probability* (3<sup>rd</sup> ed.). Oxford, UK: Oxford University Press.
- Konorski, J. (1967). *Integrative activity of the brain*. Chicago: University of Chicago Press.
- Lawrence, M. A. (2016). package ‘ez’: Easy Analysis and Visualization of Factorial Experiments. Retrieved from <https://cran.r-project.org/web/packages/ez/index.html>
- Lee, B., Gentry, R.N., Bissonette, G.B., Herman, R.J., Mallon, J.J., Bryden, D.W., Calu, D.J., Schoenbaum, G., Coutureau, E., Marchand, A.R., Khamassi, M., & Roesch, M.R. (2018). Manipulating the revision of reward value during the intertrial interval increases sign tracking and dopamine release. *PLoS Biology*, 16: e2004015.
- Lesaint, F., Sigaud, O., Flagel, S.B., Robinson, T.E., & Khamassi, M. (2014). Modelling individual differences in the form of Pavlovian conditioned approach responses: A dual learning systems approach with factored representations. *PLoS Computational Biology*, 10:e1003466.
- Mackintosh, N.J. (1974). *The psychology of animal learning*. London: Academic Press.
- Mackintosh, N.J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review*, 82, 276-298.
- Matzel, L.D., Han, Y.R., Grossman, H., Karnik, M.S., Patel, D., Scott, N., Specht, S.M., & Ghandi, C.C. (2003). Individual differences in the expression of a “general” learning ability in mice. *Journal of Neuroscience*, 23, 6423– 6433.
- Nasser, H.M., Chen, Y-W, Fiscella, K., & Calue, D.J. (2015). Individual variability in behavioral flexibility predicts sign-tracking tendency. *Frontiers in Behavioral*

*Neuroscience*, 9, 1-18.

- Patitucci, E., Nelson, A. J. D., Dwyer, D. M., & Honey, R. C. (2016). The origins of individual differences in how learning is expressed in rats: A general-process perspective. *Journal of Experimental Psychology: Animal Learning and Cognition*, 42, 313–324.
- Pavlov, I.P. (1927). *Conditioned Reflexes: An investigation of the physiological activity of the cerebral cortex*. London: Oxford University Press.
- Pearce, J.M., & Hall G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review*, 87, 532–552.
- Rescorla, R.A., & Wagner, A.R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black and W.F. Prokasy (eds) *Classical conditioning II: Current research and theory*, pp. 64-99. New York: Appleton-Century-Crofts.
- Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56, 356-374.
- Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, 22, 304-321.
- Rouder, J.N., Speckman, P.L., Sun, D., Morey, R.D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16, 225-237.
- RStudio (2015). RStudio: Integrated Development Environment for R. Boston, MA: RStudio, Inc. Retrieved from <https://rstudio.com>
- Savastano, H.I., & Miller, R.R. (1998). Time as content in Pavlovian conditioning. *Behavioural Processes*, 44, 147–162.



- Staddon, J. E. (2005). Interval timing: memory, not a clock. *Trends in Cognitive Sciences*, 9, 312–314.
- Staddon, J.E.R. & Higa, J.J. (1999) Time and memory: towards a pacemaker-free theory of interval timing. *Journal of the Experimental Analysis of Behavior*, 71, 215–251.
- Stout, S.C., & Miller, R.R. (2007). Sometimes-Competing Retrieval (SOCR): A Formalization of the Comparator Hypothesis. *Psychological Review*, 114, 759-783.
- Timberlake, W. (1993). Behavior systems and reinforcement: An integrative approach. *Journal of the Experimental Analysis of Behavior*, 60, 105-128.
- Vogel E. H., Brandon S. E., Wagner A. R. (2003). Stimulus representation in SOP: II. An application to inhibition of delay. *Behavioral Processes*, 62, 27–48.
- Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds.), *Information processing in animals: Memory mechanisms* (pp. 5-148). Hillsdale, NJ: Erlbaum.
- Wagner, A.R., & Rescorla, R.A. (1972). Inhibition in Pavlovian conditioning: Applications of a theory. In R.A. Boakes & M. S. Haliday (Eds.), *Inhibition and learning* (pp. 301-336). New York: Academic Press.
- Wilke, C. O. (2019). package ‘cowplot’: Streamlined Plot Theme and Plot Annotations for “ggplot2.” Retrieved from <https://cran.r-project.org/package=cowplot>
- Wickham, H., Henry, L., & Muller, K. (2019). package ‘dplyr’: A grammar of data manipulation. Retrieved from <https://cran.r-project.org/web/packages/dplyr/index.html>
- Wickham, H., & Henry, L. (2018). package ‘tidyr’: Easily Tidy Data with ‘spread()’ and ‘gather()’ Functions. Retrieved from <https://cran.r-project.org/web/packages/tidyr/index.html>